Segmentation via NCuts and Lossy Minimum Description Length: A Unified Approach

Mingyang Jiang, Chunxiao Li, Jufu Feng, and Liwei Wang

Key Laboratory of Machine Perception (Peking University),
MOE, Department of Machine Intelligence,
School of Electronics Engineering and Computer Science,
Peking University, Beijing 100871, P.R. China
{jiangmy,licx,fjf,wanglw}@cis.pku.edu.cn

Abstract. We investigate a fundamental problem in computer vision: unsupervised image segmentation. During the last decade, the Normalized Cuts has become very popular for image segmentation. NCuts guarantees a globally optimal solution in the continuous solution space, however, how to automatically select the number of segments for a given image is left as an open problem. Recently, the lossy minimum description length (LMDL) criterion has been proposed for segmentation of images. This criterion can adaptively determine the number of segments, however, as the optimization is combinatorial, only a suboptimal solution can be achieved by a greedy algorithm. The complementarity of both criteria motivates us to combine NCuts and LMDL into a unified fashion, to achieve a better segmentation: given the NCuts segmentations under different numbers of segments, we choose the optimal segmentation to be the one that minimizes the overall coding length, subject to a given distortion. We then develop a new way to use the coding length decrement as the similarity measure for NCuts, so that our algorithm is able to seek both the optimal NCuts solution under fixed number of segments, and the optimal LMDL solution among different numbers of segments. Extensive experiments demonstrate the effectiveness of our algorithm.

1 Introduction

A fundamental problem in computer vision is to automatically partition a natural image into regions with homogeneous texture, commonly refers to as image segmentation. Segmentation is widely accepted as a crucial function for many visual tasks such as object recognition, scene understanding and monocular inference of 3D structure. These recent vision applications have led to a renewed interest in automatic image segmentation algorithms.

In the literature, investigators have explored several important models and criteria that can lead to good image segmentation. Traditional clustering algorithms aim at extracting the statistical characteristics of the region data, such as k-means [1] and Mean Shift [2]. The graph based region merging algorithm F&H [3] attempts to partition image into regions such that the resulting segmentation is neither too coarse nor too fine. While region contours/edges contain

R. Kimmel, R. Klette, and A. Sugimoto (Eds.): ACCV 2010, Part III, LNCS 6494, pp. 213–224, 2011.

useful shape information about the saliency of the objects in the image [4], several approaches have been proposed to combine the cues of homogeneous color or texture with contours in the segmentation process [5] [6].

In recent years, much attention has been paid to spectral clustering algorithms [7], in particular the Normalized Cuts criterion [8], which provides a way of integrating global image information into the grouping process. The original NCuts criterion is concerned on the 2-way situation, which aims at partitioning image into two parts. Two recent variants extent the NCuts criterion to the k-way multi-class and multi-scale situation, known as the Multi-class NCuts [9] and Multi-scale NCuts [10]. These progress address segmentation in a k-way global optimization framework and guarantee a globally optimal solution in the relaxed continuous solution space. However, the k-way NCuts criterion can not automatically select the number of segments, k, since the objective function of k-way NCuts increases monotonically as k is varied. In many applications such as natural image segmentation, due to the diversity and complexity of image contents and semantemes, the optimal number of segments k may be different for varying images. For unsupervised segmentation, how to adaptively select the number of segments k for varying images is a fundamental open problem left in [9] and [10]. We also note that how to construct affinity matrix is another important issue in NCuts framework, which significantly influences the segmentation performance [5] [11].

More recently, an objective criterion based on the notion of lossy minimum description length (LMDL) has been proposed for evaluating segmentation of images [12]. The "optimal segmentation" is defined as the one that minimizes the number of bits needed to code the segmented data, subject to a given distortion. The most recent progress based on LMDL [13] [14] [15] have shown that this criterion is highly consistent with human segmentation of images. Preliminary success of LMDL suggests that: firstly, it is appropriate for evaluating segmentation performance objectively; secondly, the coding length serves as a reliable similarity measure between pairs of regions; last, but not the least, LMDL can adaptively determine the optimal number of segments for a given image. However, as the minimization problem is NP hard, a suboptimal solution is found by iteratively merging regions to reduce the coding length. There is no theoretical proof for the optimality of the greedy algorithm.

Although there are numerous criteria that address segmentation problem, there is little consensus on what criteria strike a best balance between objective measures that depend solely on the intrinsic statistics of imagery data and subjective measures that try to empirically mimic human perception. Some recent works such as [16] [17] focus on giving a unified perspective and evaluation procedure addressing the problem "what is a good segmentation".

Paper contributions. In this paper, we contend that, much better results can be obtained by properly combining different criteria for segmentation into a unified fashion. We propose a unified framework combining both NCuts and

LMDL criteria, to achieve an "optimal segmentation". The main novelty and the specific contributions of this paper are as follows:

1. We propose a method to automatically select the number of segments for Multi-class NCuts, using LMDL criterion. The optimal number of segments in NCuts is the one that the corresponding segmentation minimizes the overall coding length, subject to a given distortion. This perspective combines both NCuts and LMDL criteria into a unified framework, to achieve the optimal segmentation of given data.
2. We develop a new way to use the coding length decrement directly as a pairwise affinity measure, to build the affinity matrix in NCuts. This procedure allows the proposed algorithm to seek both the optimal NCuts solution under fixed number of segments and the optimal LMDL solution among different numbers of segments, thus achieves a better segmentation.
3. The experiments validate that the proposed algorithm captures the advantages of both NCuts and LMDL, thus achieves comparable or even better segmentation results compared with the state-of-the-arts.

2 Related Work

We begin by reviewing Multi-class NCuts criterion and lossy minimum description length criterion, which are closely related to our work.

2.1 k-Way Normalized Cuts Criterion

Here, we focus on the k-way Normalized Cuts [9], which means partitioning an image into k segments. Given an image I, we construct a graph $G = (V, E, W)$. Here the graph nodes V can represent either pixels or "superpixels", which is a commonly used initiation in image segmentation [18]. Suppose there are in total N nodes in the graph. Each pair of nodes is connected by a graph edge E. A weight value $W(i, j)$ represents the affinity between nodes i, j, which measures the likelihood of nodes i and j belonging to the same image segments. For a bipartition of the graph $V = V_1 \cup V_2 \cup \cdots \cup V_k, \forall V_i \cap V_j = \phi, i \neq j$, the k-way Normalized Cuts criterion is defined as:

$$\min kNcuts(V) = \frac{1}{k} \sum_{l=1}^{k} \frac{cut(V_l, V \backslash V_l)}{assoc(V_l, V)} \tag{1}$$

In the above equation, $cut(V_l, V \backslash V_l) = \sum_{i \in V_l, j \in V \backslash V_l} W(i, j)$ measures how many edge weights escape from V_l. $assoc(V_l, V) = \sum_{i \in V_l, j \in V} W(i, j)$ measures how many edge weights connects V_l.

Although directly optimizing the k-way NCuts is NP-hard, relaxing the partition indication matrix into the continuous domain turns it into a tractable continuous optimization problem and can be solved by eigenvalue decomposition of the normalized affinity matrix. This procedure is commonly known as spectral relaxing [9]. Based on the relaxed continuous solution, the final discrete solution is obtained by spectral rotation.

From Corollary 1 in [9], the k-way NCuts objective increases monotonically as k increases. This result indicates that k-way NCuts can not adaptively select the number of segments k for a given image. Consequently, *how to adaptively choose k remains an open problem.*

2.2 Lossy Minimum Description Length Criterion

In [12], Ma et.al proposed an objective measure to evaluate the quality of segmentations, which is based on the lossy minimum description length (LMDL) criterion. This criterion draws strong connection between data compression and segmentation. The optimal segmentation is defined to be the one that minimizes the number of bits needed to code the segmented data, subject to a given distortion.

First, we consider a single region V_i with m_i pixels. Based on [12], for a fixed distortion ε, the number of bits needed to code V_i under Gaussian case can be written as:

$$L(V_i) = \frac{m_i + p}{2} \log_2 \det(I + \frac{p}{\varepsilon^2} \Sigma_i) + \frac{p}{2} \log_2(1 + \frac{\mu_i^T \mu_i}{\varepsilon^2}), \tag{2}$$

where μ_i and Σ_i are the sample mean and variance of region V_i, p is the sample dimension of data.

Suppose an image I can be segmented into non-overlapping regions $V = V_1 \cup V_2 \cup \cdots \cup V_k, \forall V_i \cap V_j = \phi, i \neq j$. The LMDL criterion seeks to minimize the overall coding length of the image I:

$$min \;\; L(V_1, V_2, ..., V_k) = \sum_{i=1}^{k} [L(V_i) + m_i(-\log_2(m_i/m))] \tag{3}$$

In the above expression, m is the total number of pixels in an image, i.e., $m = m_1 + m_2 + ... + m_k$. The second term is the number of bits needed to code the membership of the m samples in the k groups (using the Huffman coding).

It is worth noticing that, once the distortion ε is fixed, the number of segments in the segmentation is *automatically determined* [12]. This completely avoids the necessity of additional interaction usually required with traditional segmentation methods, such as k-way NCuts.

However, as this minimization problem is combinatorial, all the LMDL based algorithms seek a suboptimal solution via an agglomerative way: first initialize superpixels and assume each superpixel forms its own group, then iteratively merge adjacent pair of regions that yields *the largest decrease in coding length* until the overall coding length achieves a local minima.

3 LMDL-NCuts: A Combined and Unified Criterion

We now describe our approach, which aims at combining both NCuts and LMDL into a unified fashion. We describe the general criterion below, then discuss the construction of the affinity matrix using the coding length decrement.

3.1 The Combined Criterion

From the previous section, it is clear that k-way NCuts addresses segmentation in a global optimization framework and guarantee a globally optimal solution in the continuous solution space. However, how to adaptively choose the number of segments k is left as an open problem, especially in the semantically complicated scenario such as natural image segmentation. On the other side, LMDL criterion can automatically determine the number of segments. However, the LMDL optimization is combinatorial, only local minima can be guaranteed. The complementarity of both criteria motivates us to combine both criteria into a unified fashion, to achieve a better segmentation. That is, *given the NCuts segmentations under different ks, we choose the optimal segmentation to be the one that minimizes the overall coding length, subject to a given distortion.* We refer this combined criterion to as LMDL-NCuts criterion. Note that under this criterion, the optimal number of segments is adaptively determined.

3.2 Initializing

In the original NCuts algorithm, the segmentation is directly performed on the image pixels. There are two problems for such a processing. First, each pixel will be a node in the graph so that the computational cost will be very high. Second, two pixels are connected in a graph if and only if their spatial distance is smaller than a graph connection radius G_r [10], which makes the original NCuts not able to catch the global graph topology and information. It has been investigated in [10] that larger G_r generally makes segmentation better. These problems can be solved by initializing an image with millions of pixels into a few hundred or thousand "superpixels". A superpixel is a small region that does not contain strong edges in its interior. There are several algorithms that can be used to obtain a superpixel initialization [3] [19] [20]. We have compared the three methods in the experiments and found that [20] works best for our purposes[1]. Such superpixel initialization greatly reduce the computational cost, and also, the graph connection radius constraint is no longer a necessity in our approach.

3.3 Construct Affinity Matrix

Since we build the segmentation based on the superpixel level, how to define the similarity between two superpixels is another important issue.

[1] We use the publicly available code for this method available at `http://www.cs.sfu.ca/~mori/research/superpixels/` with parameter $N_{sp} = 200$.

Since we are potentially seeking for the segmentation that can yield minimum lossy coding length, a direct way is to link the coding length with the similarity measure. In the LMDL based algorithms, the coding length decrement is used implicitly as a similarity measure between pair of regions, i.e., the decrease in the coding length essentially captures the similarity of the regions [15][2]. The larger coding length decrement means the pair of regions is more similar, so that they are merged in the LMDL based algorithms. Previous success of the LMDL based algorithms [12] [13] [15] suggests that the coding length decrement is a reliable similarity measure between regions. In our work, we directly use the coding length decrement as the affinity $W(i,j)$ between superpixels V_i and V_j. Based on Equation 2 and 3, the affinity can be written as:

$$W(i,j) = L(V_i, V_j) - L(V_i \cup V_j)$$

$$= \log_2\Big[\frac{\left|I+\frac{p}{\varepsilon^2}\Sigma_i\right|^{\frac{m_i+p}{2}}\left|I+\frac{p}{\varepsilon^2}\Sigma_j\right|^{\frac{m_j+p}{2}}}{\left|I+\frac{p}{\varepsilon^2}\Sigma\right|^{\frac{m_i+m_j+p}{2}}} \cdot \frac{(1+\frac{\mu_i^T\mu_i}{\varepsilon^2})^{\frac{p}{2}}(1+\frac{\mu_j^T\mu_j}{\varepsilon^2})^{\frac{p}{2}}}{(1+\frac{\mu^T\mu}{\varepsilon^2})^{\frac{p}{2}}} \cdot (\frac{m}{m_i})^{m_i}(\frac{m}{m_j})^{m_j}\Big]$$

Here (μ, Σ) are the sample mean and variance of region $V_i \cup V_j$, respectively. Note that $W(i,j) < 0$ means merging V_i, V_j will increase the overall coding length, which indicates that V_i, V_j are dissimilar. In this case, we simply set $W(i,j) = 0$. The calculated affinities are then normalized into the range [0,1]. The NCuts optimization is conducted based on the affinity matrix.

3.4 The Algorithm

Based on the previous discussion, we summarize the overall segmentation algorithm in Algorithm 1, which we refer to as *LMDL-NCuts Segmentation* (LNC).

Algorithm 1. LNC

1. **Input**: image data I, distortion ε;
2. **Initialization:** superpixels as graph nodes: $V := \{V_i = \{v\} \,|v \in V\}$;
3. Construct affinity matrix based on the coding length decrement;
4. **for** $k = 1 : M$ **do**
   - SegmentationResults(k) = NCuts(k);
5. Choose SegmentationResults(i) that yields the smallest overall coding length;
6. **Output**: SegmentationResults(i).

Note that for the maximum number of segments M, one can choose it to be the initial number of superpixels. However, we found that for most natural images, the number of segments larger than 40 leads to serious oversegmentation. So in our experiments, we set $M = 40$.

[2] The general spirit of a bottom-up segmentation process is to merge the "similar" regions. In the LMDL based algorithms, the pair of regions that yields the largest decrease in the coding length is merged in each iteration, which means they are "most similar" measured by the coding length similarity measure.

4 Experiments

In this section, we conduct extensive evaluation on two publicly available datasets: Berkeley Segmentation Dataset [21] and MSRC Object Recognition Dataset [22], to validate the performance of the proposed LNC algorithm. We will first describe the features used in our method. In section 4.2, we will validate that LMDL-NCuts is effective on selecting the appropriate number of segments. In section 4.3, we will discuss the effect of distortion parameter and make a close comparison with LMDL based algorithm. Both qualitative and quantitative results compared with the state-of-the-arts are listed in section 4.4.

4.1 Feature Construction

As shown in Figure 1, given an image, we convert it to the $L * a * b$ color space. In order to capture the variation of a local texture, we directly use the 7×7 cut-off window around each pixel and stack the color values inside the window into a vector form. Each texture window is smoothed by convolving with a 2D Gaussian kernel before stacking. Finally, for the ease of computation, we project the feature vectors into a D-dimensional space using PCA. We have observed that for most natural images, the first eight principal components of original feature data contain over 99% of the energy. So we set $D = 8$. The feature extraction and pre-processing are similar as in [15].

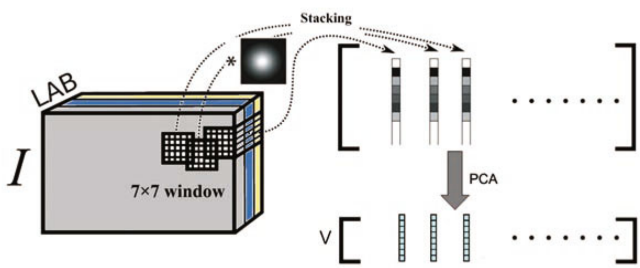

Fig. 1. Feature construction. The 7×7 windows around each pixel on the $L * a * b$ color space are convoluted with 2D Gaussian kernel, stacked into a one column vector and then use PCA.

4.2 Adaptively Select the Number of Segments

We conduct experiments on the MSRC Object Recognition Dataset [22] to validate that our LNC algorithm can adaptively select the appropriate number of segments for NCuts. MSRC dataset consists of 591 images grouped into 20 categories. Each image consists of a salient object, and the background is not so complicated so that it is visually not hard to validate the appropriate number of segments. We found that under $\varepsilon = 0.20$, the minimum description length solution provides the best visually appealing results, i.e., for most of the images, our algorithm can adaptively determine a reasonable number of segments. Some sample results are listed in Figure 2.

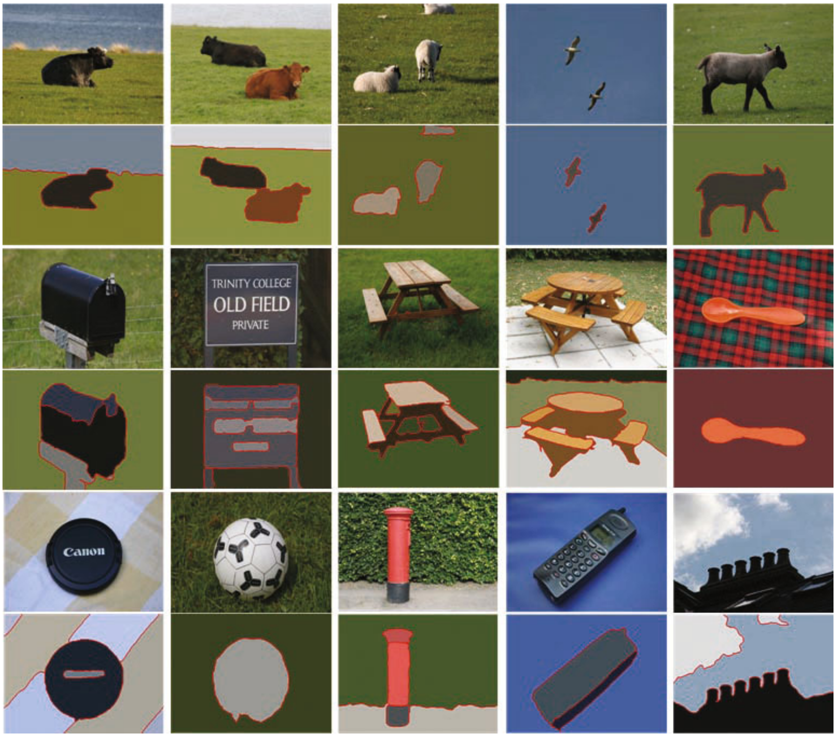

Fig. 2. (color) Qualitative results on the MSRC Object Recognition Database. For each result, the top is the original image, and the bottom is the segmentation results with each region colored by its mean color.

4.3 Comparison with LMDL Based Algorithm

We then conduct experiments on the Berkeley Segmentation Dataset (BSDS) [21]. BSDS consists of 200 training and 100 test images, and each of them has been manually segmented by a number of different subjects. Since the proposed algorithm is based on the LMDL criterion, it is worth comparing our algorithm with the LMDL based algorithms. Because we use the coding scheme proposed in the pioneer work [12], in this section, we compare LNC with the algorithm proposed in [12], namely Pairwise Steepest Descent (PSD). More comparison with other LMDL based algorithms can be found in section 4.4.

Note that the distortion ε is the only parameter in LMDL criterion. In the experiment, we compare PSD's results with our results under 4 different choices of ε: 0.10, 0.15, 0.20 and 0.25. We use three quantitative measures to evaluate the segmentation results: the average overall coding length (under given distortion), the Probability Rand Index (PRI) [23] and the Variation of Information (VoI) [24]. The objective of LMDL is to seek for the smallest coding length, so the smaller average overall coding length, the better the segmentation is. PRI and VoI aim at comparing the segmentation results with ground-truth

Table 1. Comparison between LNC algorithm and PSD algorithm. For average coding length, lower is better. For PRI, higher is better. For VoI, lower is better. Boldface indicates the better results.

	$\varepsilon = 0.10$		$\varepsilon = 0.15$		$\varepsilon = 0.20$		$\varepsilon = 0.25$	
Method/Index	LNC	PSD	LNC	PSD	LNC	PSD	LNC	PSD
avg. code length (kb)	2361	**2321**	1868	**1865**	**1562**	1573	**1331**	1364
PRI	**0.777**	0.756	**0.783**	0.758	**0.791**	0.748	**0.776**	0.724
VoI	**2.069**	2.316	**1.883**	2.062	**1.804**	1.925	**1.763**	1.896

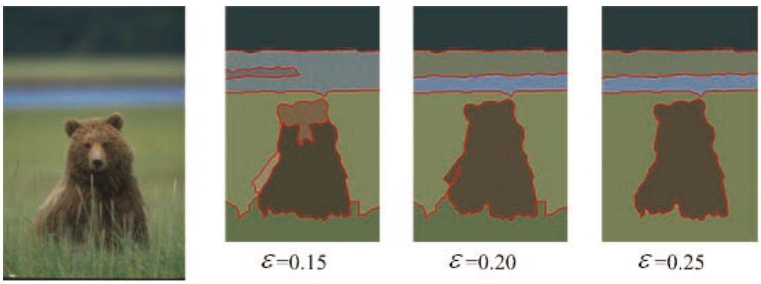

Fig. 3. (color) Segmentation results under different ε. Left: Input image. Right: Segmentation results under $\varepsilon = 0.15, 0.20, 0.25$, respectively.

results. For PRI index, the larger, the better. And for VoI, the smaller, the better. For brevity, we refer the reader to the stated references for the definition of each metric.

The results are listed in Table 1. For both PRI and VoI under all choices of ε, LNC consistently outperforms PSD. As choosing the distortion is the main difficulty in the LMDL based algorithms, here we note that the LNC algorithm is less sensitive to the choice of distortion. And, as illustrated in Figure 3, the effect of the distortion to the segmentation result is the same as in [15]: smaller choice of ε turns to over-segment images and larger ε turns to under-segment images. We also note an interesting result that LNC achieves *comparable or even smaller* average coding length compared with PSD, which is designed to directly minimize the coding length. These results suggest that the LNC algorithm captures both advantages of LMDL and NCuts, thus achieves a better segmentation.

4.4 Qualitative and Quantitative Comparison

We compare the performance of the LNC algorithm with five *publicly available* image segmentation methods: Mean-Shift (MS) [2], F&H [3], Multi-scale NCuts (MNCuts) [10], Compression-based Texture Merging (CTM) [13] and Texture and Boundary Encoding-based Segmentation (TBES) [15], on the Berkeley

Table 2. Quantitative comparison on the BSDS. Boldface indicates the best results.

Index/Method	Human	LNC	MS	FH	MNCuts	CTM	TBES
PRI (Higher is better)	0.868	**0.791**	0.772	0.770	0.742	0.742	0.787
VoI (Lower is better)	1.163	**1.804**	2.203	2.844	2.651	2.002	1.824

Fig. 4. (color) Qualitative results of LNC algorithm on the BSDS. For each result, the top is the original image, and the bottom is the segmentation with each region colored by its mean color.

Segmentation Dataset (BSDS). The performance of these five methods and that of human's, based on PRI and VOI measures, were obtained by personal communication with the authors of [15]. The user-defined parameters of these methods

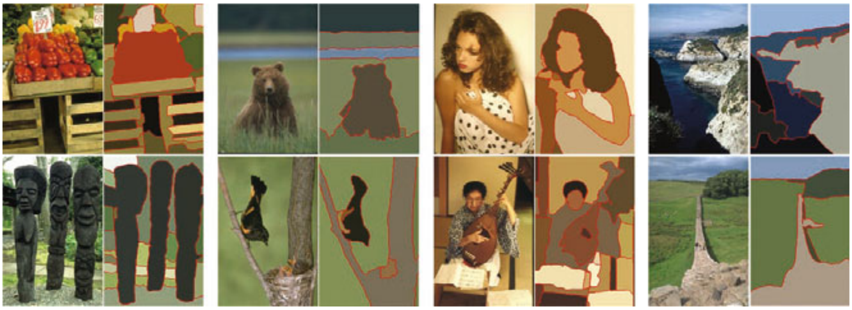

Fig. 5. (color) Qualitative results of LNC algorithm on the BSDS

have been tuned to achieve the best overall tradeoff between PRI and VoI. In particular, we report our results under $\varepsilon = 0.20$. Table 2 summarize the quantitative results on the BSDS.

Among all the algorithms in Table 2, LNC achieves the best result on both PRI and VoI. More qualitative results are shown in Figure 4 and 5.

5 Conclusion

In this paper, we have proposed a unified approach to image segmentation. It seeks both the optimal NCuts solution under fixed number of segments, and the optimal LMDL solution among segmentations under different numbers of segments. Our approach can automatically select the number of segments for NCuts, subject to a given distortion in LMDL criterion.We also develop a novel way to directly use the lossy coding length decrement as the affinity measure between superpixels, and use this affinity measure to construct affinity matrix. The experiments validate that the proposed algorithm can adaptively select the appropriate number of segments for a given image, and can yield comparable or even smaller overall coding length compared with the LMDL based algorithms. The segmentation results match well with human segmentations, compete or exceed with the best segmentation algorithms.

Acknowledgement. This work was supported by NBRPC(2011CB302400), NSFC(60635030), NSFC(61075003) and NSFC(60775005).

References

1. Duda, R., Hart, P., Stork, D.: Pattern Classification. John Wiley and Sons, Chichester (2001) 0-471-05669-3
2. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. PAMI 24, 603–619 (2002)
3. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. IJCV 59, 167–181 (2004)
4. Ren, X., Fowlkes, C., Malik, J.: Learning probabilistic models for contour completion in natural images. IJCV 77, 47–63 (2008)

5. Malik, J., Belongie, S., Leung, T.K., Shi, J.: Contour and texture analysis for image segmentation. IJCV 43, 7–27 (2001)
6. Zhu, S.C., Tu, Z.W.: Image segmentation by data-driven markov chain monte carlo. PAMI II, 131–138 (2002)
7. Chung, F.R.K.: Spectral graph theory. Regional Conference Series in Mathematics, vol. 92, pp. 1–212. American Mathematical Society, Providence (1997)
8. Shi, J., Malik, J.: Normalized cuts and image segmentation. PAMI 22, 888–905 (2000)
9. Yu, S.X., Shi, J.B.: Multiclass spectral clustering. In: ICCV, pp. 313–319 (2003)
10. Cour, T., Benezit, F., Shi, J.B.: Spectral segmentation with multiscale graph decomposition. In: CVPR, vol. II, pp. 1124–1131 (2005)
11. Kim, T.H., Lee, K.M., Lee, S.U.: Learning full pairwise affinities for spectral segmentation. In: CVPR (2010)
12. Ma, Y., Derksen, H., Hong, W., Wright, J.: Segmentation of multivariate mixed data via lossy data coding and compression. PAMI 29, 1546–1562 (2007)
13. Yang, A.Y., Wright, J., Ma, Y., Sastry, S.S.: Unsupervised segmentation of natural images via lossy data compression. CVIU 110, 212–225 (2008)
14. Rao, S.R., Mobahi, H., Yang, A.Y., Sastry, S.S., Ma, Y.: Natural image segmentation with adaptive texture and boundary encoding. In: Zha, H., Taniguchi, R.-i., Maybank, S. (eds.) ACCV 2009. LNCS, vol. 5994, pp. 135–146. Springer, Heidelberg (2010)
15. Mobahi, H., Rao, S.R., Yang, A.Y., Sastry, S.S., Ma, Y.: Segmentation of natural images by texture and boundary compression. In: arXiv:1006.3679v1 (2010)
16. Bagon, S., Boiman, O., Irani, M.: What is a good image segment? A unified approach to segment extraction. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 30–44. Springer, Heidelberg (2008)
17. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: From contours to regions: An empirical evaluation. In: CVPR, pp. 2294–2301 (2009)
18. Ren, X.F., Malik, J.: Learning a classification model for segmentation. In: ICCV, pp. 10–17 (2003)
19. Ren, X.F., Fowlkes, C.C., Malik, J.: Scale-invariant contour completion using conditional random fields. In: ICCV, vol. II, pp. 1214–1221 (2005)
20. Mori, G.: Guiding model search using segmentation. In: ICCV (2005)
21. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: ICCV, vol. 2, pp. 416–423 (2001)
22. Shotton, J., Winn, J., Rother, C., Criminisi, A.: Textonboost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 1–15. Springer, Heidelberg (2006)
23. Rand, W.M.: Objective criteria for the evaluation of clustering methods. American Statistical Association Journal 66, 846–850 (1971)
24. Meila, M.: Comparing clusterings: An axiomatic view. In: ICML (2005)